Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis

D. Dong,[1] P. Fang,[2] Y. Bock,[2] F. Webb,[1] L. Prawirodirdjo,[2] S. Kedar,[1] and P. Jamason[2]

Received 28 April 2005; revised 15 November 2005; accepted 9 December 2005; published 9 March 2006.

[1] Spatial filtering is an effective way to improve the precision of coordinate time series for regional GPS networks by reducing so-called common mode errors, thereby providing better resolution for detecting weak or transient deformation signals. The commonly used approach to regional filtering assumes that the common mode error is spatially uniform, which is a good approximation for networks of hundreds of kilometers extent, but breaks down as the spatial extent increases. A more rigorous approach should remove the assumption of spatially uniform distribution and let the data themselves reveal the spatial distribution of the common mode error. The principal component analysis (PCA) and the Karhunen-Loeve expansion (KLE) both decompose network time series into a set of temporally varying modes and their spatial responses. Therefore they provide a mathematical framework to perform spatiotemporal filtering. We apply the combination of PCA and KLE to daily station coordinate time series of the Southern California Integrated GPS Network (SCIGN) for the period 2000 to 2004. We demonstrate that spatially and temporally correlated common mode errors are the dominant error source in daily GPS solutions. The spatial characteristics of the common mode errors are close to uniform for all east, north, and vertical components, which implies a very long wavelength source for the common mode errors, compared to the spatial extent of the GPS network in southern California. Furthermore, the common mode errors exhibit temporally nonrandom patterns.

Citation: Dong, D., P. Fang, Y. Bock, F. Webb, L. Prawirodirdjo, S. Kedar, and P. Jamason (2006), Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis, *J. Geophys. Res.*, *111*, B03405, doi:10.1029/2005JB003806.

1. Introduction

[2] Continuous GPS (CGPS) networks consisting of hundreds of global stations and thousands of regional stations provide a unifying framework to monitor surface deformation patterns from global to regional to local scales. Studies using CGPS measurements have revealed deformation patterns over a range of spatial scales from global plate motion [*Larson et al.*, 1997; *Sella et al.*, 2002; *Prawirodirdjo and Bock*, 2004] and geocenter motion [*Blewitt et al.*, 2001; *Dong et al.*, 2003] to regional crustal deformation at plate boundaries [*Shen et al.*, 1997; *Murray et al.*, 1998; *Calais*, 1999; *Miyazaki and Heki*, 2001; *Beavan et al.*, 2002; *Márquez-Azúa and Demets*, 2003; *Hammond and Thatcher*, 2004] to local coseismic deformation [*Bock et al.*, 1993; *Hudnut et al.*, 1996; *Simons et al.*, 2002]. Quantification of the deformation patterns are useful for investigating geophysical processes with various temporal variations from postglacial rebound [*Johansson et al.*, 2002] to postseismic deformation and slow earthquakes [*Dragert et al.*, 2001; *Miller et al.*, 2002; *Hudnut et al.*, 2002] to seasonal deformation [*vanDam et al.*, 1994; *Hatanaka et al.*, 2001; *Dong et al.*, 2002] to transient motions [*Melbourne and Webb*, 2002; *Melbourne et al.*, 2002; *Larson et al.*, 2003; *Bock et al.*, 2004; *Ohtani et al.*, 2004]. Accompanying the efforts of scientists to extract more information from CGPS, various filtering techniques have been developed to suppress or eliminate various noise sources in the position time series. In regional network analysis, the so-called common mode error (CME) is one of the major spatially correlated error sources in the CGPS solutions, which is mitigated through a technique commonly referred to as regional filtering.

[3] Regional spatial filtering of CGPS position time series was first introduced by *Wdowinski et al.* [1997] to improve the resolution of coseismic and postseismic displacements for the 1992 M_w 7.6 Landers earthquake in southern California. This approach removes a common mode bias from each coordinate component, computed by "stacking" the position residuals of the stations and estimating and removing the mean value. This procedure is equivalent to a three-parameter Helmert transformation of the origin of the network (i.e., a 3-D translation). The "filtered" time series have a significantly lower root mean square error thereby improving the resolution of the observing system for

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA.
[2]Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, La Jolla, California, USA.

0148-0227/06/2005JB003806$09.00

detecting weak and transient signals [*Wdowinski et al.*, 1997; *Smith et al.*, 2004]. Error analysis of CGPS position time series shows that regional filtering reduces the amplitudes of both the white noise and flicker noise components by a factor of 2–3 [*Williams et al.*, 2004]. Several GPS data analysis centers have adopted this or similar approaches to regional filtering in their routine work and provide both unfiltered and filtered time series for public use (e.g., http://sopac.ucsd.edu/cgi-bin/refinedJavaTimeseries.cgi; http://reason.scign.org; http://sideshow.jpl.nasa.gov/mbh/series.html). The regional filtering technique has been widely used as part of the study of various tectonics processes [*Wdowinski et al.*, 1997; *Calais*, 1999; *Scherneck et al.*, 2000; *Parker*, 2001; *Miller et al.*, 2001a, 2001b; *Miyazaki et al.*, 2003; *Márquez-Azúa and Demets*, 2003; *Wdowinski et al.*, 2004; *Smith et al.*, 2004].

[4] Although widely used, regional spatial filtering is performed in a somewhat empirical fashion. There are several interesting questions that should be further investigated. For example, what is the physical source of the CME source? What signals are contained in the common mode bias? Should the spatial distribution of the CME be uniform over the whole region? How can this methodology be extended to larger regional networks? In this paper, we show that common mode spatial filtering can be understood in the context of spatiotemporal filtering by principal component analysis (PCA) and the Karhunen-Loeve expansion (KLE). We will demonstrate that the combination of PCA and KLE provides a more solid numerical framework for addressing these basic questions, and improving the extraction of signal and systematic error from GPS regional position time series.

2. Regional Filtering Approach

[5] Daily position time series from the Southern California Integrated GPS Network (SCIGN) are used in this paper, specifically the Scripps Orbit and Permanent Array Center's (SOPAC) most recent operational time series concatenated onto the reanalysis by *Nikolaidis* [2002] of global IGS and SCIGN data from January 1991 to January 2002. The SOPAC analysis uses a distributed processing method described by *Zhang* [1996]. A global analysis divides the IGS stations into a number of interleaved as well as regionally clustered subnetworks, with three common stations between subnetworks. The a priori positions of the so-called IGS core stations are assigned tight constraints, and each subnetwork is processed with the GAMIT software [*King and Bock*, 2005]. The estimated parameters include site positions, troposphere delays and gradients, orbits and Earth orientation parameters (EOP). The model terms applied include solid earth tide, ocean tide, antenna phase center variation, and pole tide. The elevation cutoff is set to 10 degrees, and there is no ambiguity resolution. There is also an elevation-dependent weighting applied in the process using postfit residuals. Once individual subnetwork solutions are obtained, they are combined with the GLOBK software [*Herring*, 2005]. Data from the regional stations are analyzed in a similar manner, but with ambiguity resolution and tightly constraining the IGS orbit and EOP parameters at the GAMIT step. The parameter adjustments and their respective covariance matrices for the global and regional solutions are output as loosely constrained solutions. The global and regional solutions are combined using the GLOBK software on a daily basis, and aligned by a weighted seven-parameter transformation to the ITRF2000 reference frame [*Altamimi et al.*, 2002]. More details of the data processing are given by *Nikolaidis* [2002]. The set of individually estimated daily positions then make up the "raw" position time series for the global and regional stations.

[6] The raw regional position time series contain both signal and noise with temporal and spatial correlation. The noise characteristics of CGPS time series have been investigated in the time domain [*Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams et al.*, 2004] and in the space domain [*Wdowinski et al.*, 1997; *Nikolaidis*, 2002]. It is standard practice to model the time series on a component-by-component (north, east, vertical) basis with terms for constant episodic offsets (due to coseismic deformation, or instrumental change), linear trends (from tectonic motion and long-period processes, such as postglacial rebound; could have multiple rates), nonlinear variations (from postseismic deformation, magma intrusion; usually modeled by simple analytical functions, such as exponential or logarithmic decay), and seasonal (annual and semiannual) terms (from surface mass loading, such as atmosphere, ocean, groundwater, and underground aquifer). The resulting position time series are called in this paper "residual" time series. See *Nikolaidis* [2002] for more details on this procedure. The residual time series contain various systematic (from network common to site-dependent) and random errors, as well as unmodeled signals.

[7] The regional spatial filtering approach, referred to as "stacking" in this paper, calculates a common mode bias in the detrended and demeaned residual position time series, component by component. As presented by *Nikolaidis* [2002], the weighted CME from a set of s representative stations is computed for days $i = 1, \ldots, m$ by

$$\varepsilon(t_i) = \frac{\sum_{k=1}^{S} \left(\nu_k(t_i)/\sigma_{i,k}^2 \right)}{\sum_{k=1}^{S} \left(1/\sigma_{i,k}^2 \right)} \tag{1}$$

where $\nu_k(t_i)$ is the residual value of the kth station at epoch i and $\sigma_{i,k}$ is its standard error. For each station component, the filtered position $x(t)$ is derived simply by subtracting the common mode value from the observed component value $x^0(t)$,

$$x(t) = x^0(t) - \varepsilon(t) \tag{2}$$

This approach works well for regional networks such as SCIGN but has certain limitations when extended over larger regions such as North American plate [*Márquez-Azúa and Demets*, 2003], since the assumption of spatial uniformity breaks down and the common mode bias becomes progressively smaller.

[8] In this paper, we explore a more general spatiotemporal filtering approach, allowing nonuniform spatial response of the network stations to a CME source, but still assuming a uniform temporal function across the network. For a regional network daily station coordinate time series

with n stations and spanning m days, the ($m \times n$) real-valued matrix $\mathbf{X}(t_i, x_j)$ ($i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$) (without loss of generality, we assume $m \geq n$) is constructed. In $\mathbf{X}$, each column contains the detrended and demeaned coordinate values for a single geodetic component (north, east, or vertical) from a single station in the network, and the rows contain geodetic component values for all stations at a given epoch. The element of covariance matrix $\mathbf{B}$ used in our presentation is defined as

$$b_{ij} = \frac{1}{m-1}\sum_{k=1}^{m} X(t_k, x_i)X(t_k, x_j) \tag{3}$$

[9] The ($n \times n$) symmetric matrix $\mathbf{B}$ can be decomposed as

$$B = V\Lambda V^T \tag{4}$$

where eigenvector matrix $\mathbf{V^T}$ is a ($n \times n$) matrix with orthonormal rows, the Λ matrix has k nonzero diagonal eigenvalues $\{\lambda_k\}$ ($n \geq k$). In most cases with real geodetic data, the rank of matrix $\mathbf{B}$ is usually full ($k = n$). From linear algebra, any matrix of rank n can be expanded by an n orthonormal vector basis. Thus we choose the orthonormal function basis $\mathbf{V}$ to expand the data matrix $\mathbf{X}(t_i, x_j)$

$$X(t_i, x_j) = \sum_{k=1}^{n} a_k(t_i)v_k(x_j) \tag{5}$$

where $a_k(t)$ is derived by

$$a_k(t_i) = \sum_{j=1}^{n} X(t_i, x_j)v_k(x_j) \tag{6}$$

Such decomposition is called empirical orthogonal function (EOF) analysis [*Menke*, 1984], also known as principal component analysis (PCA) [*Preisendorfer*, 1988]. The $a_k(t)$ is called the kth principal component (PC) of matrix $\mathbf{X}$, and $\mathbf{v}_k(x)$ is its corresponding eigenvector. The principal components represent the temporal variations, and the eigenvectors represent the corresponding spatial responses to the principal components. If we arrange the eigenvectors so that the eigenvalues are in descending order, the first few PCs represent the biggest contributors to the variance of the network residual time series, usually related to the common source time function; The higher-order PCs are usually related to local or individual site effects. Such a PCA approach was successfully applied to the decomposition of geodetic data for the study of interseismic deformation [*Savage*, 1988, 1995; *Scherneck et al.*, 2000; *Parker*, 2001] and regional filtering [*Johansson et al.*, 2002].

[10] When the covariance matrix $\mathbf{B}$ is normalized by the variance vector σ, the resultant matrix becomes the correlation matrix $\mathbf{C}$, i.e., $c_{ij} = b_{ij}/(\sigma_i\,\sigma_j)$, where σ is defined as

$$\sigma_j = \sqrt{b_{jj}} = \sqrt{\frac{1}{m-1}\sum_{k=1}^{m}\left[X(t_k, x_j)\right]^2} \tag{7}$$

[11] The correlation matrix can also be decomposed as

$$C = W\Lambda_c W^T \tag{8}$$

where the eigenvectors $\mathbf{W}$ are not the same as $\mathbf{V}$ in (4), and the $\Lambda_\mathbf{c}$ matrix is also different from the Λ matrix in (4). The data matrix $\mathbf{X}(t_i, x_j)$ can also use the orthonormal vector basis $\mathbf{W}$ to perform the expansion, and the decomposition formulas are similar to (5) and (6).

$$X(t_i, x_j) = \sum_{k=1}^{n} a_k(t_i)w_k(x_j) \tag{9}$$

$$a_k(t_i) = \sum_{j=1}^{n} X(t_i, x_j)w_k(x_j) \tag{10}$$

Such an expansion is called Karhunen-Loeve expansion (KLE), which was originally used in pattern recognition [*Fukunaga*, 1990] and later introduced to geophysical and geodetic problems [*Rundle et al.*, 2000; *Tiampo et al.*, 2004].

[12] The only difference between PCA and KLE is that the former uses the covariance matrix $\mathbf{B}$ and the latter uses the correlation matrix $\mathbf{C}$ to compute the orthonormal vector basis. If we rescale the original time series by the square roots of their variance reciprocals, the KLE analysis of the original time series is equivalent to the PCA analysis of the rescaled time series. Here the rescaled data matrix $\tilde{X}$ is defined as

$$\tilde{X}(t_k, x_j) = \frac{X(t_k, x_j)}{\sigma_j} \tag{11}$$

where σ is defined in (7). That is, the PCA decomposition arranges the eigenvectors based on their contributions to the network time series variance. Similarly, the KLE analysis provides the eigenvectors based on their contributions to the rescaled time series variances. It is easy to show that the rescaled data matrix suppresses the time series with larger variance, so that every rescaled time series has the same normalized variance.

3. Regional Filtering of GPS Station Position Residual Time Series

3.1. Residual Time Series and Preliminary Analysis

[13] We use a 5-year span (2000–2004) of daily coordinate solutions from the SOPAC analysis of SCIGN data. We set the criteria of sufficient data availability as 72%, which restricts our analysis to those 152 stations that span more than 3.6 years. For each epoch, we set 5% as the effective epoch criterion. That means if at this epoch there are only less than eight stations having solutions, this epoch will be discarded. For the 5-year span SOPAC daily solutions, 4 days are discarded. For each coordinate time series, we estimate a constant offset, trend, and annual and semiannual terms as described by *Nikolaidis* [2002]. Then we subtract these terms from the coordinate time series to form the residual time series. The coordinate jumps due to antenna changes are also estimated and are removed from the original time series. Our time series starts ten weeks after the 16 October 1999 M_w 7.1 Hector Mine earthquake, and we model the effects of postseismic deformation [*Pollitz et al.*, 2000, 2001; *Fialko*, 2004] using the

exponential decay model from *Nikolaidis* [2002]. The postseismic effects are also removed from the time series.

[14] The north, east and vertical residual station position time series are analyzed independently to form three data matrices. The daily coordinate solutions with formal uncertainties larger than the chosen thresholds of 50, 50, and 100 mm for east, north, and vertical components, respectively, are discarded. When the residuals exceed the thresholds of 100, 100, and 300 mm for east, north, and vertical components, respectively, they are considered as outliers and are discarded. We fill time gaps shorter than 2 days with a three-point Lagrangian interpolation. For larger gaps we first set the values at these missing days using the spatially averaged values derived from all effective sites on these days. Then we construct the covariance and correlation matrices to perform PCA and KLE analysis, embedded in the QOCA software [*Dong et al.*, 1998; http://gipsy.jpl.nasa.gov/qoca], to get principal components $a_k(t)$ and eigenvectors $\mathbf{v_k}(x)$ and $\mathbf{w_k}(x)$. Taking only the first three principal modes and using equations (5) and (9), we perform PCA and KLE analysis iteratively until convergence to refill in these missing points. Here we set the convergence criteria as the average change less than 10^{-6} of the missing point value from the previous iteration. For our data, the convergent results are reached in less than 3 iterations.

[15] Next, we perform both PCA and KLE analysis separately on the continuous (no gaps) residual coordinate matrix. We construct the covariance and correlation matrices **B** and **C** and decompose them to get eigenvalues and eigenvectors following standard procedures (reducing **B** and **C** to a symmetric tridiagonal matrix by Householder transformation, deriving eigenvalues and eigenvectors from the tridiagonal matrix using QL algorithm). We then compute the principal components $a_k(t)$ by equation (6) (for PCA) and equation (10) (for KLE). The principal component $a_k(t)$ and eigenvector $\mathbf{v_k}(x)$ and $\mathbf{w_k}(x)$ together are considered as "mode k", where $a_k(t)$ represents a temporal signature of the mode and $\mathbf{v_k}(x)$ and $\mathbf{w_k}(x)$ represents the spatial footprint of the mode. To make the comparison more instructive, we divide each eigenvector by its maximum (absolute value) element, so that the response of this "maximum" element is always 100% and all spatial response values are always in the range (−100%, 100%). We call such an eigenvector a normalized eigenvector. We multiply the corresponding principal component by the normalization factor. Such a principal component is referred to as a scaled PC.

3.2. Sensitivity Study

[16] To understand the consequences of the PCA and KLE treatments on real data, we perform two simulation tests. To mimic the real GPS time series, we simply take the first four principal components of the KLE analysis of the 2000–2004 daily vertical coordinate time series of 149 SCIGN stations (stations MHMS, SACY and FXHS excluded) as the simulation mode series (Figure 1). The first mode is chosen as CME and the other three modes as local systematic effects. We assume the regional network consists of four stations. Note that the empirical orthogonal functions are related to the real residual time series of the 149 stations. For the four stations, the four principal components are no longer totally uncorrelated. The time series at each station are generated by summation of their spatial responses

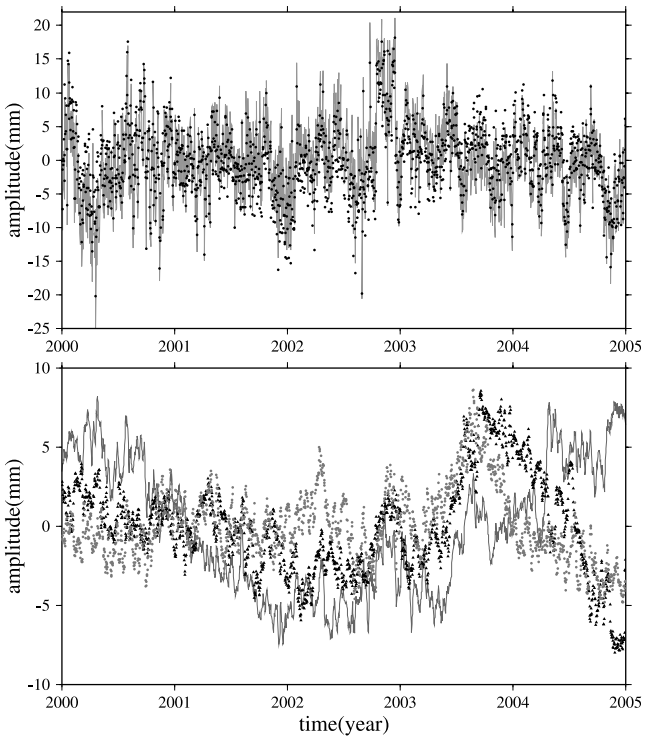

Figure 1. Simulation time series. (top) Mode 1 time series (solid line) used as the CME for both simulations 1 and 2. The black dots represent the recovered CME series in test 2 by KLE approach. (bottom) Local mode time series used in simulation 2. The solid line is mode 2. The black triangle is mode 3. The grey dot is mode 4.

to CME and other local systematic effects, as well as random noises. Here the spatial response is defined as the amplitude ratio of the CME at this station to the raw simulation CME series. We want to test if the CME series can be extracted and if the spatial responses at each station can be recovered.

[17] In the first simulation, we assume the first mode in Figure 1 as the CME series for four stations. Their spatial responses are 100%, 70%, 80%, and 90%, respectively. The time series of all the four stations also contain Gaussian noise (zero mean, root of variance 2 mm), which are at the same level as the CME. Both approaches recover the CME series quite well. The recovered spatial responses of the four stations are 100%, 71%, 81%, 92% for PCA approach, and 100%, 98%, 99%, 100% for KLE approach, respectively. In this case (only CME and random noises exist), the PCA approach retains correct spatial responses for the four stations. However, the KLE approach obtains biased spatial responses. The reason is that the KLE approach incorrectly interprets the time series with stronger spatial responses as "noisy" series, and the rescaling reduces the responses of these time series.

[18] In the second simulation, we still assign the first mode as the CME series for all the four stations with 100%, 70%, 80%, and 90% spatial responses, respectively. Meanwhile, station 1 contains mode 2 time series (500% amplitude of Figure 1), station 2 contains Gaussian noises with 2 mm root of variance, station 3 contains mode 3 time series (100%

amplitude of Figure 1), and station 4 contains mode 4 time series (50% amplitude of Figure 1). In this test the CME is no longer the dominant component, and stations 1, 3, and 4 have strong local systematic effects or random noise, in particular for station 1. We do not expect that the simulated CME can be completely recovered from the time series, because both PCA and KLE decompositions only construct the empirical orthogonal functions and the locally affected time series (modes 2, 3, 4) are not totally uncorrelated with the CME series. In PCA analysis, the mode 2 series at station 1 becomes the first principal mode. The CME series become the second mode with the spatial responses of −11%, 74%, 92% and 100% for stations 1, 2, 3, and 4, respectively. In the KLE analysis, the first principal mode is CME with the spatial responses of 21%, 96%, 96% and 100% for stations 1, 2, 3, and 4, respectively. In EOF decomposition, each successive principal mode is the best fit of the data in $\mathbf{X}$ with contributions of previous modes removed [*Savage*, 1995]. Thus the PCA analysis always constructs the principal component, which best deduces the time series variance, no matter whether the component is due to common mode errors or local effects, or a combination of the two. That is the reason why PCA identifies the mode 2 time series as the first PC. On the other hand, any local effects increase the time series variance. The KLE approach rescales the time series to get normalized variances, and hence reduces the influences of local effects (also reduces the spatial responses of CME in these time series). When there is a coherent signal in the network time series with multiple stations involved, such a rescaling property effectively suppresses the local effects and makes the coherent signal much more significant.

[19] In the second simulation test the CME time series can be recovered by both the KLE (as first PC) and PCA (as second PC) approaches, but both have small low-frequency deviations due to the interference from local mode effects. We plot the KLE approach recovered CME time series in Figure 1 for comparison. This demonstrates that the KLE approach reduces the influences of local effects more effectively than PCA and can recover the CME even in the presence of strong local effects (however the resultant spatial responses of CME are biased). On the other hand, the PCA approach is preferred when the time series contain CME, random noises and weak local effects. In our simulation test, when the local effects (mode 2 series) at station 1 are reduced to 100% amplitude level, the PCA approach is able to recover the CME and spatial responses of the four stations satisfactorily. This test represents an extreme case (strong local effects and only four stations). When the network station number increases, the contribution of coherent common signal (or CME) to the variance becomes more and more overwhelming, so that the threat from uncorrelated local effects is greatly mitigated.

[20] The simulations tell us that the KLE approach has the capability of extracting coherent signals even from noisy time series with strong local effects, but the recovered spatial responses are not always accurate. Hence the KLE approach is most suitable for noisy data or for the applications, such as pattern recognition, which focus on detecting hidden coherent signals, and where the recovery of the spatial responses is not critical. The PCA approach, however, is able to recover both coherent signals and the correct spatial responses but it is biased when the local effects are dominant. Hence the PCA approach is most suitable for data where the coherent signals are dominant and the local effects have been effectively suppressed. Since the goals of this paper are to identify and extract both the CME temporal variations and spatial responses in the residual time series, we utilize both the PCA and KLE approaches. The KLE approach is used to identify the CME and other coherent patterns and to prevent the PCA approach from going astray when there are strong local effects. The PCA approach is used to obtain the CME time series and spatial responses in the final stage when the CME is identified and the local effects are sufficiently suppressed.

3.3. Identify Stations With Strong Local Effects

[21] For the residual time series with 152 SCIGN stations, both the PCA and KLE results obtain similar scaled PC time functions and nearly uniformly distributed normalized spatial eigenvectors for the first PC of the north component. For the east component, the first PC time series from PCA and KLE are similar. However, the first PC of PCA results show three stations (MHMS, SACY, and FXHS) with abnormally large responses. In particular, the response of MHMS is about 80% larger than the responses of most stations. The remaining stations have nearly uniform spatial responses. On the other hand, the spatial responses of the KLE results show nearly uniform distribution for all stations because the responses at stations MHMS, SACY and FXHS are suppressed. Since the normalization factor is the maximum response of the network stations, the east amplitudes of the first scaled PC series from KLE results are about 60% of that from PCA results. Meanwhile, the normalized spatial responses at most stations of the PCA results are about 60% of that of the KLE results. There are two possible mechanisms to explain such a difference. One possibility is that the abnormally large responses at the three stations are real, and the KLE approach down weighs the three time series and makes their apparent spatial responses similar to that of other stations. If such an explanation is true, we must accept a special nonuniform spatial response distribution: nearly uniform distribution at most stations and abnormally large responses at the three stations. Another possibility is that strong local effects at the three stations contaminate their first east PC responses. High-order east PCs of the PCA results reveal that strong local effects exist at the three stations. The east residual time series of the three stations also show abnormal variations with several big data gaps. Furthermore, all the three stations are located in the northwestern rim of the Santa Ana basin, where the aquifer activities are strongest [*Bawden et al.*, 2001; *Watson et al.*, 2002; *Argus et al.*, 2005]. We do not attempt to model the local effects at the three stations and just simply discard them from our station list.

[22] After removing the three stations, the PCA and KLE analyses show much more consistent results for the remaining 149 stations. It is interesting that that the KLE results are nearly the same with or without the three discarded stations. Figures 2, 3, and 4 show the PCA results of the spatial and temporal patterns of the first two principal modes for north, east, and vertical residual coordinate time series, respectively. The KLE results are very similar to the final PCA results (see next section), and so we do not display them in this paper. The

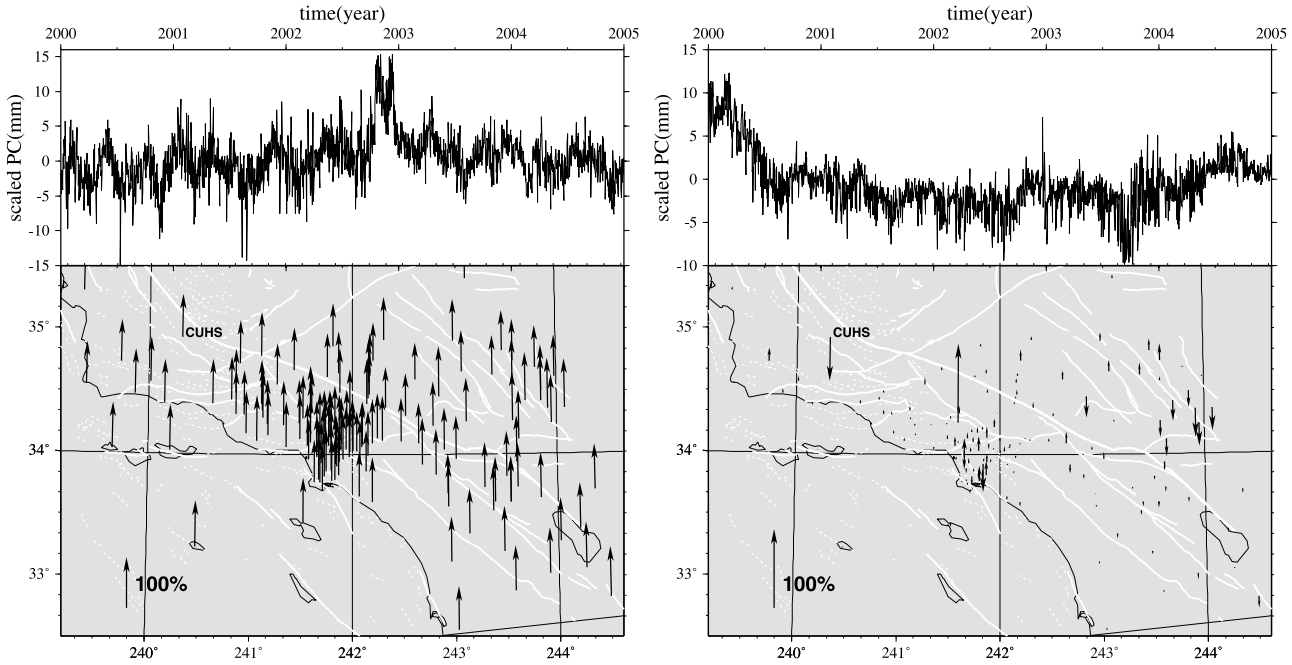

Figure 2. East component of PCA solution. (left) (top) First scaled PC (solid line) and (bottom) its normalized spatial eigenvectors. The grey dot series (offset by −0.012 m) in Figure 2 (top) represent the east displacement time series at USC1 caused by atmospheric mass loading. The arrows represent the element values of the normalized eigenvectors (not the displacement directions). The up arrows represent positive responses to the scaled PC; the down arrows represent negative responses to the scaled PC. Station name CUHS is printed next to its arrow. (right) (top) Second scaled PC and (bottom) its normalized spatial eigenvectors. The arrows are defined as in Figure 2 (left).

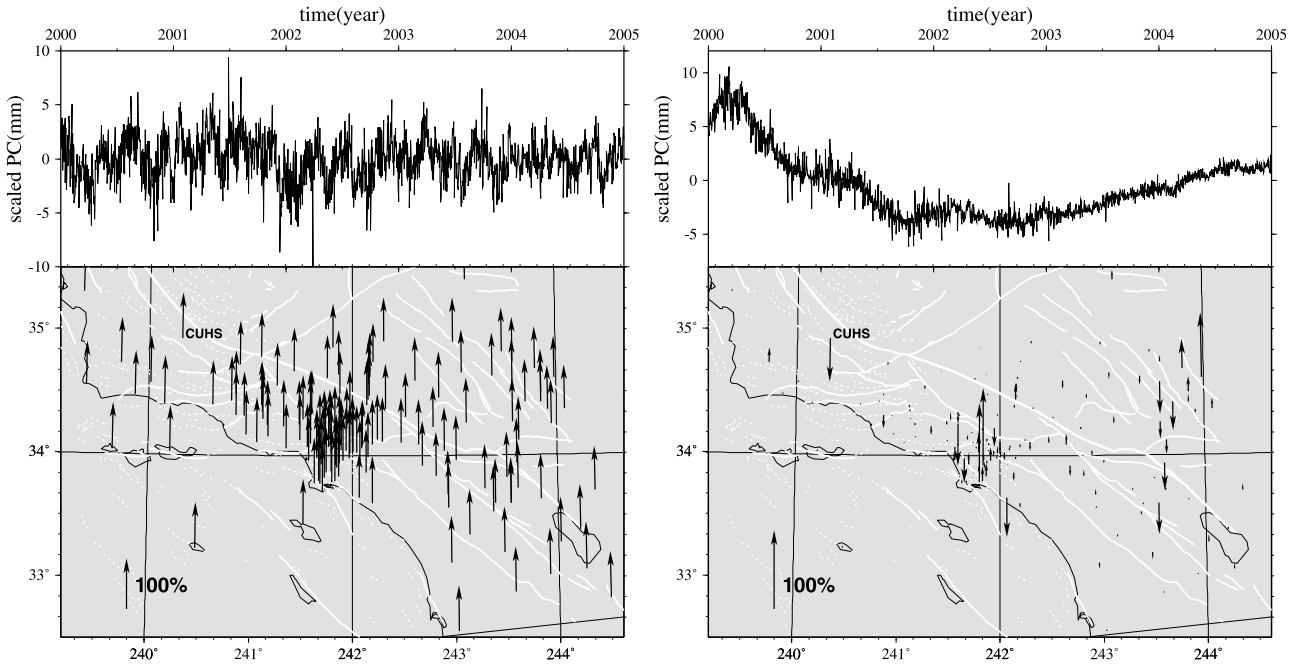

Figure 3. North component of PCA solution. The arrows are defined as in Figure 2. (left) (top) First scaled PC (solid line) and (bottom) its normalized spatial eigenvectors. The grey dot series (offset by −0.008 m) in Figure 3 (top) represent the east displacement time series at USC1 caused by atmospheric mass loading. (right) (top) Second scaled PC and (bottom) its normalized spatial eigenvectors.

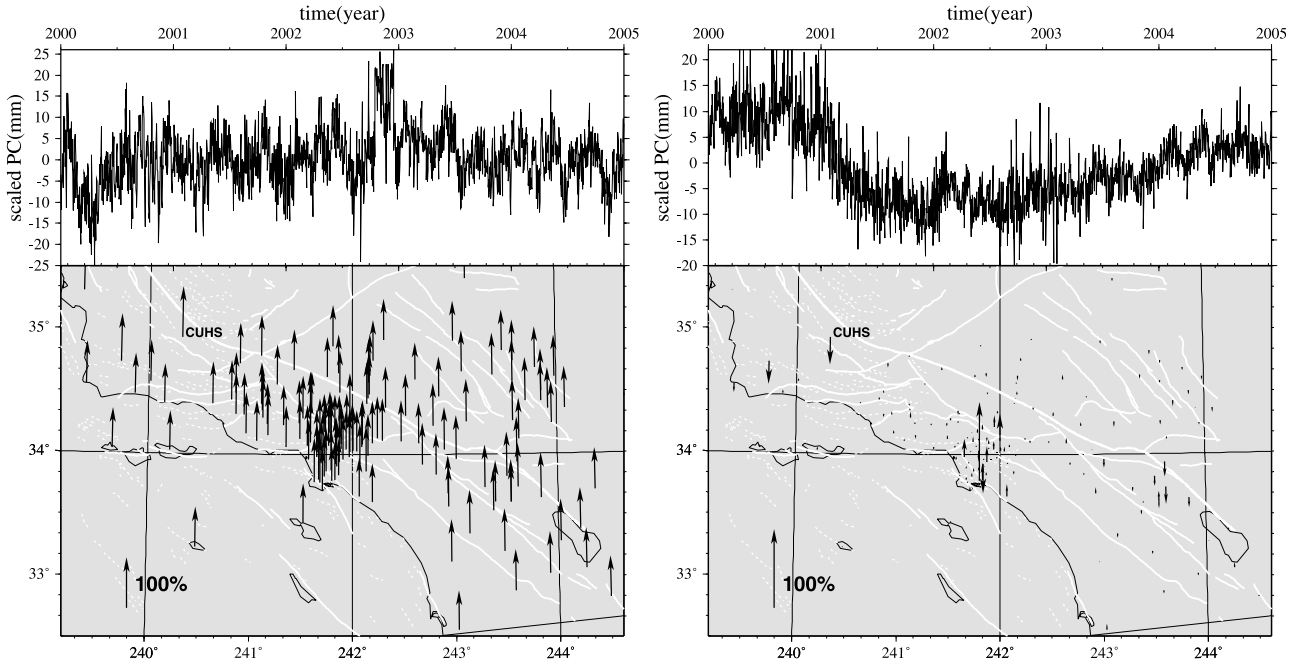

Figure 4. Vertical component of PCA solution. The arrows are defined as in Figure 2. (left) (top) First scaled PC (solid line) and (bottom) its normalized spatial eigenvectors. The grey dot series (offset by −0.020 m) in Figure 4 (top) represent the east displacement time series at USC1 caused by atmospheric mass loading. (right) (top) Second scaled PC and (bottom) its normalized spatial eigenvectors.

scaled first PC time series as well as the normalized spatial eigenvectors (all east, north and vertical components) of the PCA and KLE results are similar except the vertical PCA spatial response at station CUHS is about 10% larger than the responses of all other stations (Figure 4, left). To check if the slightly abnormal vertical response at station CUHS is real or from local effects, we examine the patterns of the second- and high-order PC.

[23] The temporal variations in the second PC are dominated by the long-period (>1.2 years) components (Figures 2, 3, and 4, right). Their spatial distribution indicates that less than 20% of the stations have significant responses, and not all are of the same sign. Here we define "significant" as the (absolute value) normalized response either more than 25% or more than twice the (absolute value) average normalized response. The average normalized responses (absolute values) of the 149 stations are 10%, 10%, and 6% for east, north, and vertical components, respectively. The stations with significant responses in the east and north series are located in the northwestern rim of the Santa Ana basin and the Mojave Desert near the epicentral regions of the Landers and Hector Mine earthquakes. These may be related to aquifer water volume variations and unmodeled postseismic deformation. Moderate responses in the vertical series are seen for stations in the Salton Trough and northwestern rim of the Santa Ana basin. Both areas have strong aquifer activities. This mode also looks like an east-west tilt motion (see Figure 4, right). Across longitude 243°, the stations on the east and west sides move in opposite vertical directions. In the western Transverse Ranges area, there is one isolated station (CUHS) with significant responses to all east, north and vertical second PCs. We are unable to identify potential geophysical processes at the station. An examination of the third mode's spatial eigenvectors indicates that only a small portion of the stations have modest amplitudes with some significance. Their temporal variations exhibit from high-frequency irregular fluctuations to interannual variations. These high-order PCs are probably the mixture of unmodeled signals, local effects and noise. The second PCs of the KLE results confirm that station CUHS has significant spatial responses for all east, north and vertical components. Comparing both first and second PCs, we postulate that the station CUHS has strong local effects from unknown local sources. Therefore we also remove the station CUHS from the next step of our analysis, leaving a total of 148 stations.

3.4. Pattern Analysis of the PCA and KLE Results

[24] Using the residual time series of the 148 stations, we repeat the PCA and KLE analysis. The eigenvalues ($\mathbf{\Lambda}$ in (4) or $\mathbf{\Lambda}_c$ in (8)) of the east, north, and vertical covariance or correlation matrices represent the share of each PC mode in the total covariance or correlation, and their cumulated shares are displayed in Figure 5. The first PC eigenvalues from KLE analysis show higher percentages than that from PCA analysis, indicating that KLE approach enhances the coherent common modes, while suppressing the local effects effectively. It is clear that the residual time series do have significant correlated modes. The first modes represent 89%, 76%, and 78% (PCA analysis) and 90%, 81%, and 80% (KLE analysis) of the total eigenvalues for the east, north, and vertical components, respectively.

[25] The first mode PCs (temporal functions and spatial eigenvectors) of the east, north, and vertical components of the PCA analysis for the 148 stations are displayed in Figure 6, 7, and 8. Since the stations with strong local

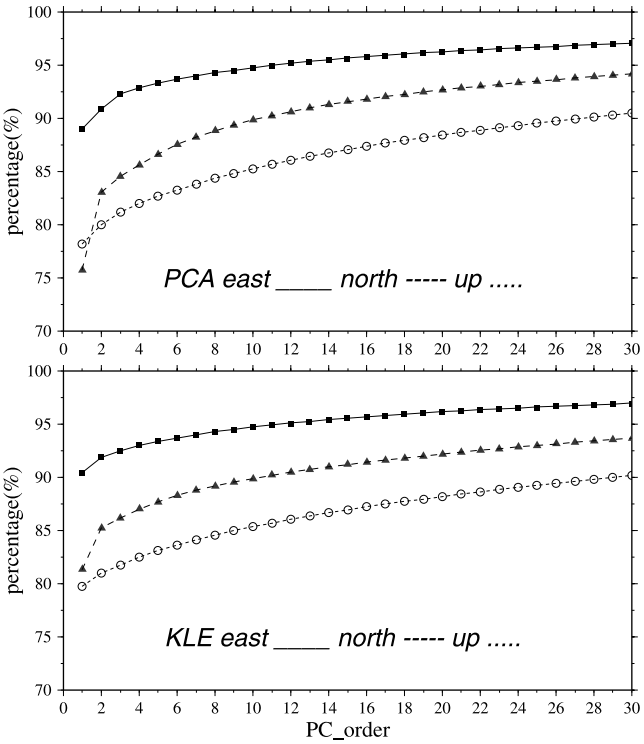

Figure 5. Cumulative percentage of PC eigenvalues. For visual clarity, we only plot the eigenvalues of the first 30 PCs. (top) PCA result. (bottom) KLE result.

effects have been identified and excluded, the KLE results are very similar to the PCA results. The first PC spatial eigenvectors of all 148 GPS stations have the same sign, with average normalized amplitudes (absolute values) of 86%, 87%, and 86% from PCA analysis and 96%, 92%, and 93% from KLE analysis for east, north, and vertical components, respectively. The minimum normalized amplitudes are 69%, 76%, and 71% from PCA analysis and 50%, 47%, and 51% from KLE analysis for east, north, and vertical components, respectively. The temporal variations of the east, north, and vertical coordinates over the 5-year time span are not purely random, they display some systematic structure. Spectral analysis shows that for both PCA and KLE results there are significant features with 3–4 months period in all three components, and significant approximately 14-day periods in the vertical component. Figure 9 shows the spectrum results of the PCA analysis. The 14-day period spectrum might be related to the ocean tide model (Mf tide, semimonthly lunar tide with period of 13.7 days) used in the analysis. The GAMIT software performs ocean tide corrections using Scherneck's model [*Scherneck*, 1991], which includes the Mf tide correction. One possibility is that the model is still not good enough for the Mf tide, in particular near the southern California coastal area. Another possibility is model error in the diurnal O1 tide and semidiurnal M2 tide, whose alias periods are also around 14 days [*Penna and Stewart*, 2003]. Though the most likely origin of this signal may stem from imperfect tidal correction, other potential sources cannot be ruled out. We leave the problem for future investigation.

[26] Since the trends and seasonal terms have been subtracted in the residual time series, there is no apparent tectonic process common to all of southern California stations that could explain the first PC. The potential geophysical candidates for the first PC series are mass loading from atmosphere, oceans and groundwater [*Mangiarotti et al.*, 2001; *vanDam et al.*, 2001; *Dong et al.*, 2002]. Aquifer water volume variations can cause significant site motion [*Bawden et al.*, 2001; *Watson et al.*, 2002; *Argus et al.*, 2005] but only on a more local scale. In order to verify if the atmospheric mass loading plays any role, we performed a spot check on the first PC series of one typical station, USC1. Since the dominant spatial wavelength of the deformation caused by atmospheric mass loading is expected at the 10^3 km level, we calculated the atmospheric loading caused displacements at station USC1 as a representative of the network as a whole. The normalized eigenvectors for USC1 are 84%, 91% and 84% in PCA analysis and 99%, 97%, and 91% in KLE analysis for the east, north, and vertical components, respectively, and so are representative of the average spatial responses of the network stations. The 6-hour sampling National Center for Environmental Prediction (NCEP) reanalysis surface pressure data from 2000 to 2004 are used to calculate the daily displacements at USC1. The inverted barometer (IB) model is implemented for the oceanic response to surface pressure variations. The Green's function approach [*Farrell*, 1972; *vanDam and Wahr*, 1987] is adopted in the calculation. To be consistent with the residual time series, we remove a bias, trend, and annual and semiannual terms from the calculated displacement time series due to atmospheric mass loading to obtain the residual time series. The residual displacement time series at USC1 are also plotted in Figures 6, 7, and 8 (top). As can be seen in Figures 6, 7, and 8, the amplitudes of the atmospheric loading caused motions are too small and the correlations with the PC time series are too poor to be the source of the first PC. Thus the dominant motions of the first PC series are likely nonphysical and related to mismodeled satellite orbits and EOP, and reference frame limitations [*Wdowinski et al.*, 1997].

4. Regional Filtering Using PCA and KLE Approach

4.1. Comparisons With the Stacking Approach

[27] Here we first examine the nature of the filtered time series derived by the stacking approach. Equation (1) indicates that the common mode bias is a spatially weighted average and therefore it is network-dependent. Equation (2) assumes that the CME has spatially uniform distribution and therefore it is not network-dependent. When the CME is computed independently for each day, the existence of daily CME is implied. In this sense, the stacking procedure can be considered as a spatiotemporal filter. The assumption of spatial uniform distribution of CME over the network limits the application of the stacking approach. As the size of the regional network increases, we expect that the calculated magnitude of the daily CME will be reduced. In the extreme case of a global network, the concept of CME loses its meaning, as the common mode bias is equivalent to a global translation. The successful application of stacking approach to regional networks such as SCIGN indicates that the regional network size is smaller than the common mode wavelength. After gaining better knowledge of the charac-

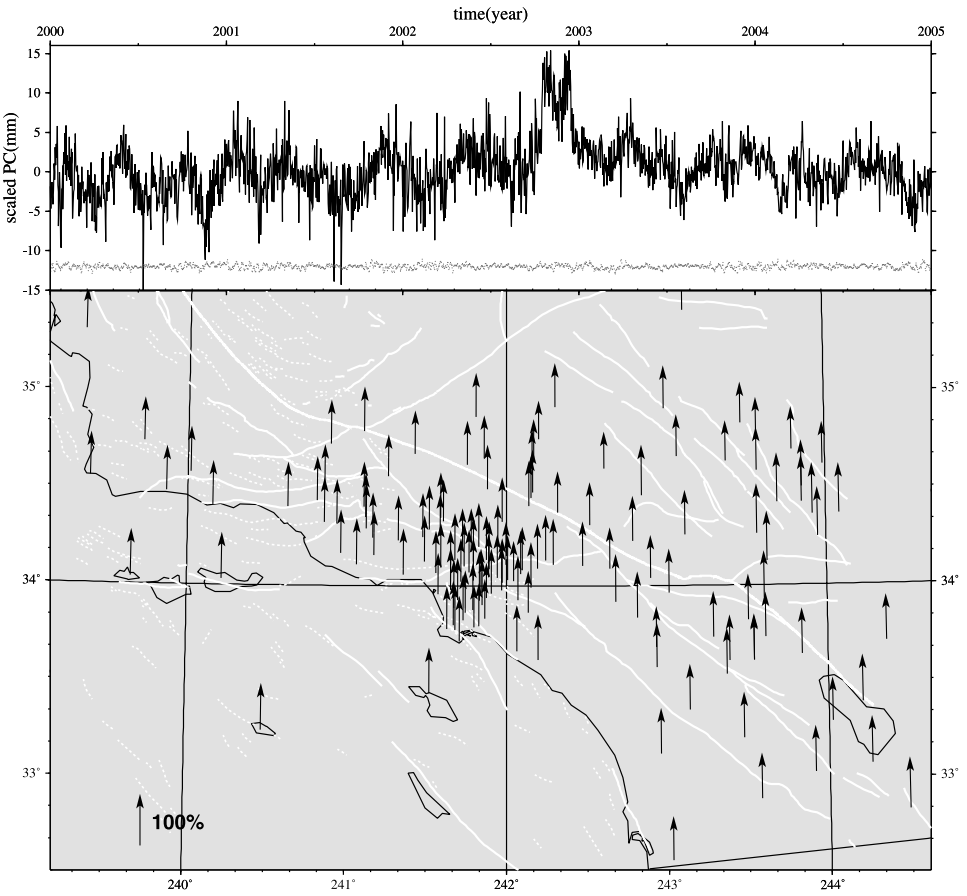

Figure 6. First east PC component of PCA solution. The arrows are defined as in Figure 2. (top) Scaled PC (solid line) and (bottom) its normalized spatial eigenvectors.

teristics of the CME, especially the wavelengths of various systematic errors, it is desirable to redefine the CME as a function of the wavelength.

[28] Our spatiotemporal filtering approach (PCA and KLE) retains the assumption that the CME is both spatially and temporally correlated for all regional network stations, but implements the CME filtering in a more general way. In the space domain, our spatiotemporal filtering approach removes the uniform distribution constraint and allows the network data themselves to reveal the spatial distribution of the CME. In the time domain, the stacking approach assumes the network residual time series contain only CME and random errors. The PCA and KLE approaches assume that the network residual time series contain CME and various local effects and random noises, whose spatiotemporal patterns are orthogonal to the spatiotemporal patterns of CME. It is apparent that the spatiotemporal patterns of random noise are orthogonal to uniformly distributed CME, so that the stacking approach is just a special case of the general spatiotemporal filtering approach. The actual network residual time series contain many modes from network common to local common and site-dependent. The stacking approach simply performs a spatially weighted average on a daily basis and essentially separates the network data into two modes: uniformly distributed and randomly distributed. Thus the effects of other local modes will partly leak into the uniformly distributed mode and partly enter the randomly distributed mode. The EOF analysis takes the entire network time series into account and decomposes the time series into various spatial and temporal coherent orthogonal modes. Although these local effects in general are not completely orthogonal to CME, they are close to orthogonal if the network common mode signals are removed in advance and the individual local effects are not strongly correlated. Thus the EOF decomposition is able to provide better separation than the stacking approach. Furthermore, the spatial distribution of the CME in the real data is not available with the stacking approach, as it already assumes uniform distribution in advance. On the other hand, the PCA and KLE approaches are able to display the spatiotemporal distribution of each principal mode for more detailed investigation.

[29] Next, we discuss the weighting of the PCA, KLE and the stacking approaches. Equation (6) indicates that the PCA derived CME can also be considered as a "weighted" mean using the same spatial eigenvector $\mathbf{v}(x_k)$ as a weighting function for all days, while the stacking approach uses different $1/\sigma_{i,k}^2$ as the weighting function from day to day. The $\sigma_{i,\mathrm{k}}$ is a daily standard error at epoch i and station k, which reflects the solution formal uncertainty, usually is related to random observation errors but not necessarily related to CME. For example the station with strong local effects or weak CME response still can have small standard error (hence large weighting). When the real data indeed

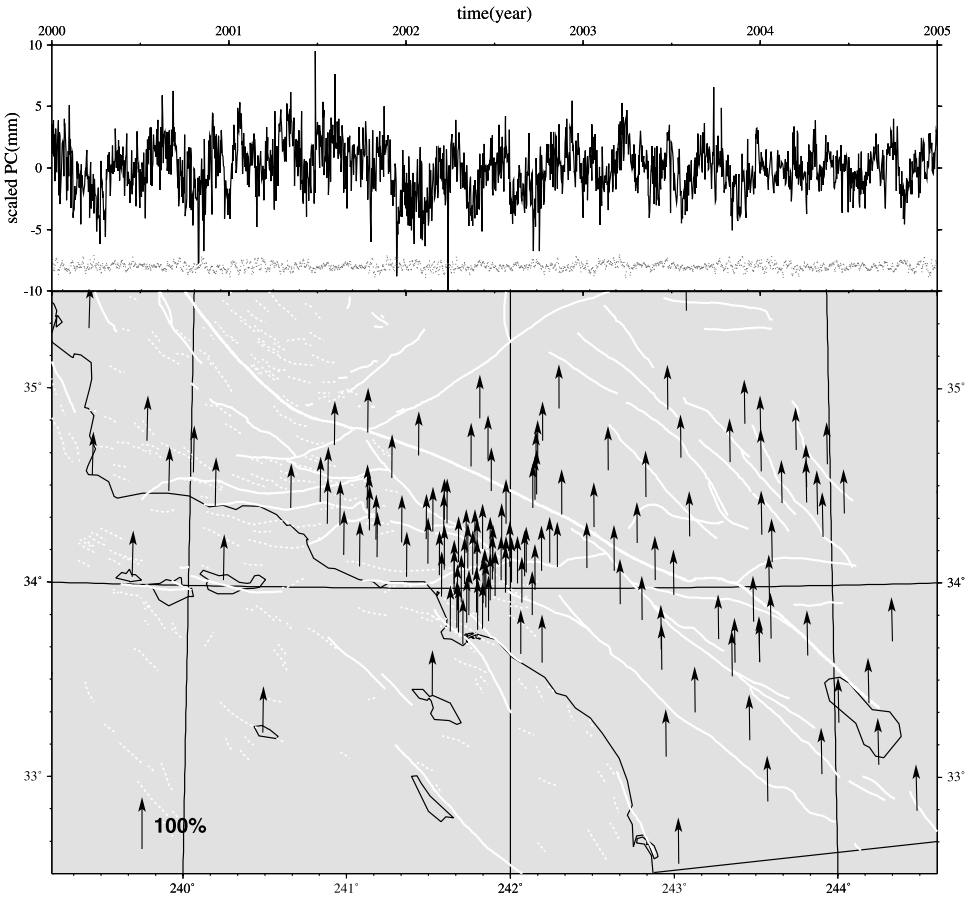

Figure 7. First north PC component of PCA solution. The arrows are defined as in Figure 2. (top) Scaled PC (solid line) and (bottom) its normalized spatial eigenvectors.

have only spatially uniform distributed CME and random errors, this weighting function of the stacking approach is meaningful. When the real CME distribution violates the uniform assumption or there are significant local effects in the data, the weighting function of the stacking approach is not optimal. The $\mathbf{v}(x_k)$ in equation (6) reflects the spatial correlation among the network stations derived from the time series of the entire network, which is related to a network coherent pattern such as CME, assuming that any strong local effects have been suppressed effectively. However, our current PCA weighting formulation ignores the formal uncertainties, so that it is vulnerable to the perturbations from weak or noisy data. Thus the PCA approach is most suitable for relatively homogeneous data sets with weak or noisy data excluded in advance. The weight function we use for the KLE approach is a combination of PCA weighting and time series variance, so that it resists noise and local effects better but could also suppress the pattern at stations with strong responses. How to define an optimal weighting function to perform spatiotemporal filtering is still an open question for future investigation.

[30] We further illustrate the possible different outcomes from three typically encountered cases. 1. A station has very weak CME, so that its temporal variations differ significantly from that of other network stations. 2. A station has normal CME, but it also has local effects partially correlated with CME. 3. A station has normal CME, but deviates for a short (few days) period of time (perhaps due to some transient deformation or local error). In the first case, the poor correlation between this station and the rest of the network is embodied in the covariance matrix **B** with small off-diagonal terms (absolute values) related to this station. The CME related eigenvector (derived from (4)) has a very small response element related to this station. Therefore this station basically does not contaminate the common mode PC (calculated by (6)). Also, a small CME correction is imposed on this station due to a small response in the eigenvectors. In the second case, PCA constructs the orthogonal eigenvectors to minimize the variance. Because of the correlation between CME and the local effects at the station, the station's corresponding element in the CME eigenvector is enlarged. The impact on the common mode PC is small because most stations do not have the similar correlated local effects (otherwise this local effect is no longer "local", it should be another CME). The CME corrections on all other network stations are basically unbiased except this station, which is overcorrected. In the third case, since the covariance matrix is constructed using the entire time series, this few day aberration will not significantly affect the overall correlation in PCA analysis. In all three cases, the CME derived by stacking will be biased if this station is included in the list of representative stations. That is,

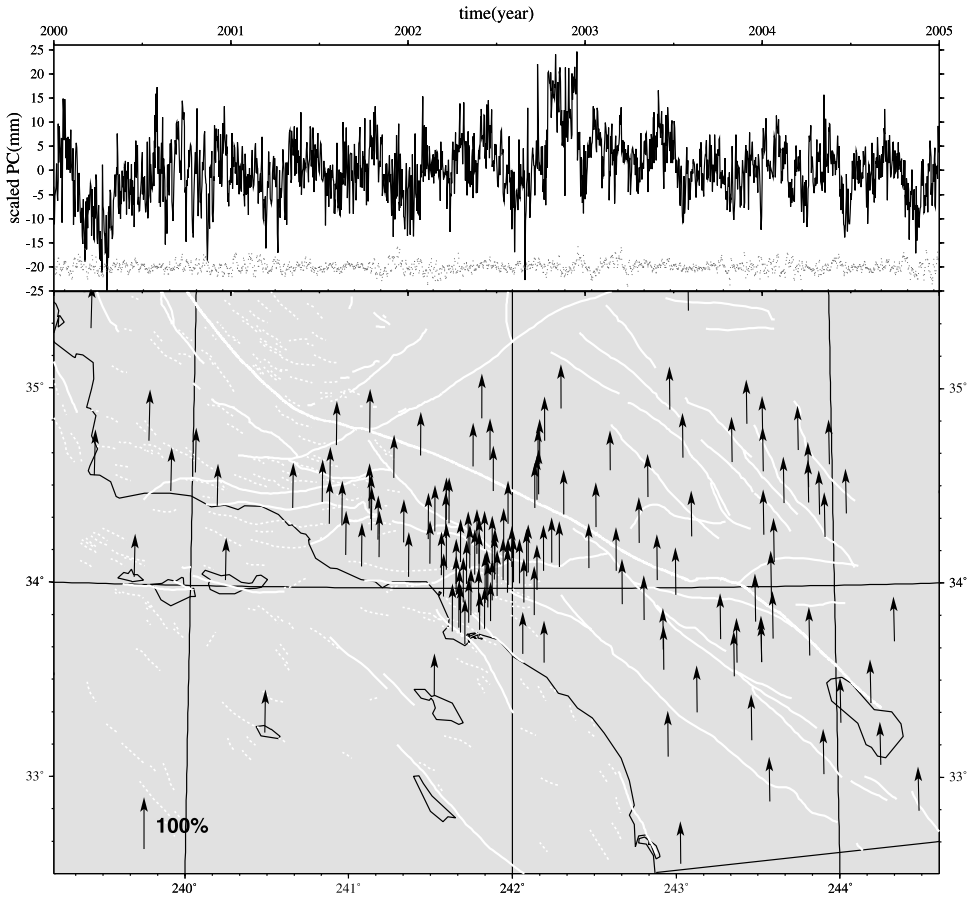

Figure 8. First vertical PC component of PCA solution. The arrows are defined as in Figure 2. (top) Scaled PC (solid line) and (bottom) its normalized spatial eigenvectors.

the choice of the representative stations is subjective. In the first case, this poorly correlated station gets an erroneous CME correction, so that its filtered time series is degraded (larger RMS) compared with its unfiltered time series. These simple (yet common) examples indicate that our spatiotemporal filtering approach is not sensitive to a single poorly correlated (in space or time) station. It avoids a potential bias due to overly relying on a set of selected stations for the common mode calculation.

4.2. Define the CME From PCA Analysis

[31] Although the CME and regional spatial filtering are two widely used terms, the definition of ''common mode error'' has not been rigorously and quantitatively defined. Another question raised by our proposed approach is: how to identify which mode is the common mode? We notice that there is no absolute boundary between the network common mode and subnetwork common modes, and there is no generally accepted consensus. In this paper, we tentatively treat the mode as the common mode if most sites (more than 50%) have significant normalized responses (larger than 25%), and the eigenvalues of this mode exceed 1% of the summation of all eigenvalues. According to this criterion, only the top few PC modes are potential candidates for the common mode, because the high-order modes are usually related to a few stations and presumably reflect local effects. In general, the common mode PC contains both errors and unmodeled common signals. The PC can be expressed as

$$a_k(t_i) = a_k^s(t_i) + a_k^e(t_i) \tag{12}$$

where $a_k^s(t_i)$ and $a_k^e(t_i)$ denote unmodeled common signals and common errors, respectively. If the signals and errors are separable, only the common errors represent CME. The CME from PCA is defined by

$$\varepsilon_j(t_i) = \sum_{k=1}^{p} a_k^e(t_i) v_k(x_j) \tag{13}$$

where p is the identified common mode PC number. For KLE analysis, the CME is derived using the same equation (13) where the $\mathbf{v_k}(x)$ is replaced by $\mathbf{w_k}(x)$.

[32] For our SCIGN residual time series, based on the discussion of the previous section, only the first modes ($p = 1$) of the east, north, and vertical components satisfy our criteria. In this paper, we simply treat the first PC mode as CME, although this mode still contains unmodeled signals such as caused by nonseasonal atmospheric mass loading, which is much smaller than the PC (see Figures 6, 7, and 8). We cannot identify any geophysical process,

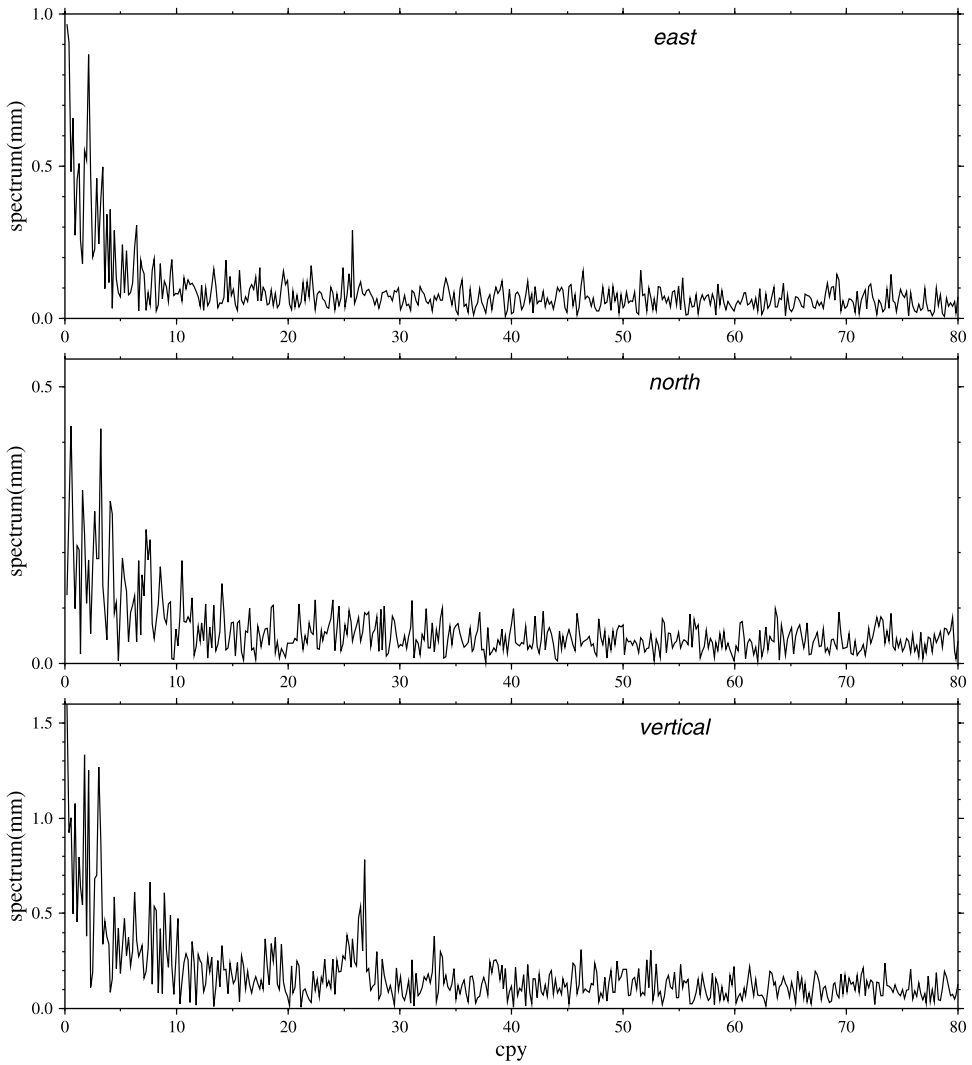

Figure 9. Power spectrum of the first PC time series from PCA analysis. For visual clarity, we only plot the spectrum to cpy (cycles per year) = 80. (top) East component. (middle) North component. (bottom) Vertical component.

which would make a significant contribution to the first PC mode.

4.3. Comparisons

[33] Figure 10 shows a comparison of the CME time series (years 2000–2004) derived by the stacking and PCA approaches. To be consistent with the stacking results, the PCA derived CME time series are scaled by the spatially averaged response instead of the maximum response. The stacking results are derived from the SOPAC operational analysis. The selected stations for stacking are not the same for each day, and are based on the solution quality for each day. In the 2000–2004 period, the maximum selected stations are 60 and the minimum selected stations are 17. Figure 10 indicates very good agreement for all east, north and vertical components, which confirms that the SCIGN daily solution time series contain the CME, and the time series of the common mode biases derived by the stacking is a good representation of the CME. There is one day (2002.6671) when the stacking CME time series of the north components have significant larger correction than the PCA CME value. The reason is that some of the chosen representative stations had abnormal local effects on this day. This anomaly is not present in the PC time series, since the PCA approach is less sensitive to a single abnormal station.

[34] It is interesting to compare the scatters of the filtered time series between PCA and stacking approaches. The weighting function (spatial eigenvector) of the PCA is derived from the entire network time series, and the same spatial weighting function is used for each day, while the weighting function of the stacking approach comes from the standard error of each daily solution. Thus the comparisons of overall mean scatters and daily scatter time series between the two approaches will give us some clues about the nature of the CME, for example, whether the CME is the dominant error in the residual network time series? Are the spatial responses of the CME close to uniform? Which weighting function reduces the scatters more effectively?

[35] The filtered daily mean scatter time series of the two approaches are displayed in Figure 11. The daily

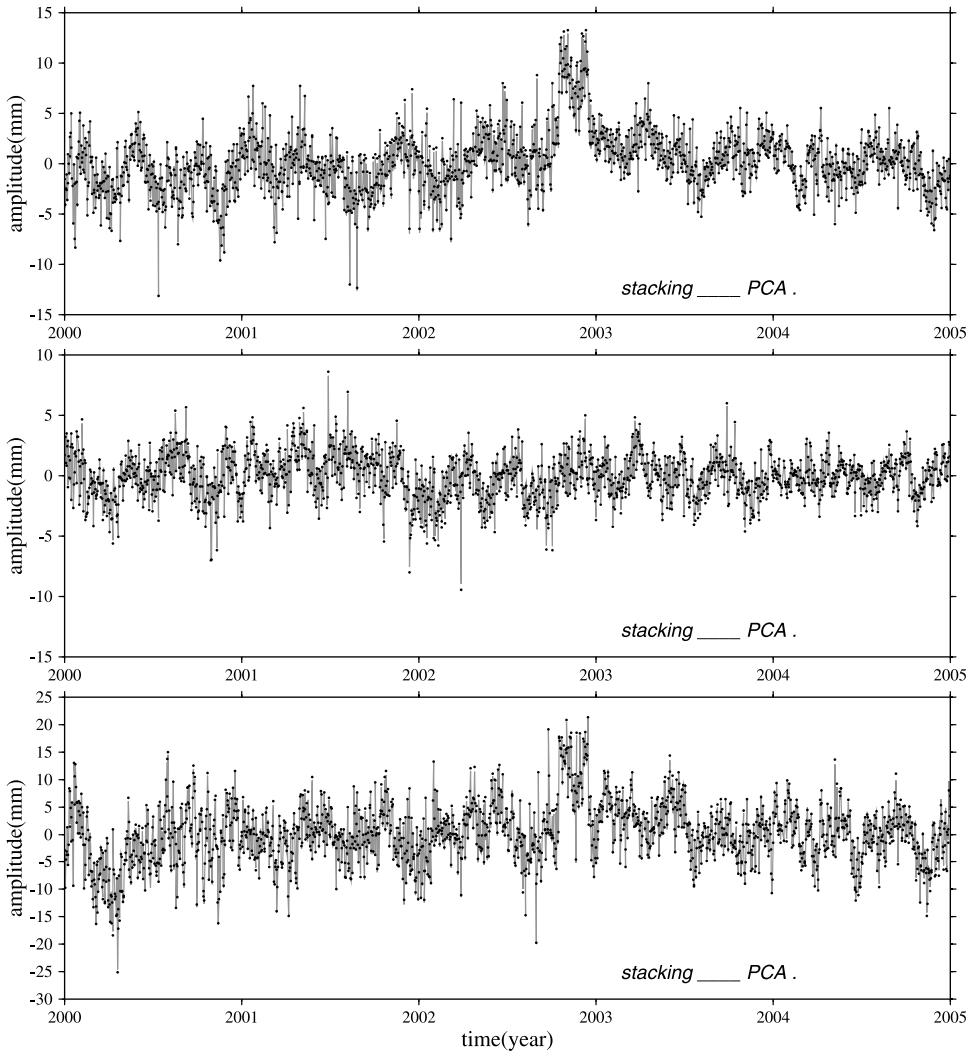

Figure 10. Comparison of the common mode time function. The solid line is the stacking solution and the triangle is the PCA solution. (top) East component. (middle) North component. (bottom) Vertical component.

scatters in Figure 11 are calculated using the sum of absolute residual values divided by station number (L1 norm scatters) to mitigate the influences from abnormal residuals, which are usually related to local anomalies. Since the daily spatial mean CME time series from PCA analysis are very close to the stacking results (Figure 10), the slightly smaller scatters of the PCA results (Figure 11) indicate that the slightly nonuniform spatial responses (eigenvector) from the PCA analysis fit the data better. There are few days when both stacking and PCA approaches have large scatters, implying some abnormal local effects on these days. For example, on day 2002.6671 there are only 22 regional stations including several poorly correlated stations in the northwestern rim of the Santa Ana basin. The weak regional solutions on that day are dominated by other local effects, which are not consistent with the CME from the entire time series. Table 1 lists the overall mean L1 and L2 norm scatters calculated of the 5-year scatter time series for east, north and vertical components from the stacking, PCA and KLE approaches. Where the daily L1 norm scatter is defined by the sum of absolute residual values divided by station number, the daily L2 norm scatter is the square root of the sum of residual squares divided by station number. The Table 1 indicates that all the three approaches are able to reduce the scatters significantly for the SCIGN daily solutions. Our regional filtering (PCA results) reduces the scatter power (calculated using L2 norm scatters) of the unfiltered residual time series by 87.6%, 73.1% and 75.5% for east, north and vertical components, respectively. The stacking approach reduces the scatter power by 87.0%, 72.8%, and 74.9% for east, north, and vertical components, respectively.

[36] It should be mentioned that regional filtering not only improves the resolution of the observing system for detecting weak and transient signals, but also uncovers other contributors to the CGPS error budget. For example, the vertical daily scatters of the filtered time series (Figure 11) clearly demonstrate that the network residual scatters are larger during the summer than that during the winter. This is likely related to error sources that are

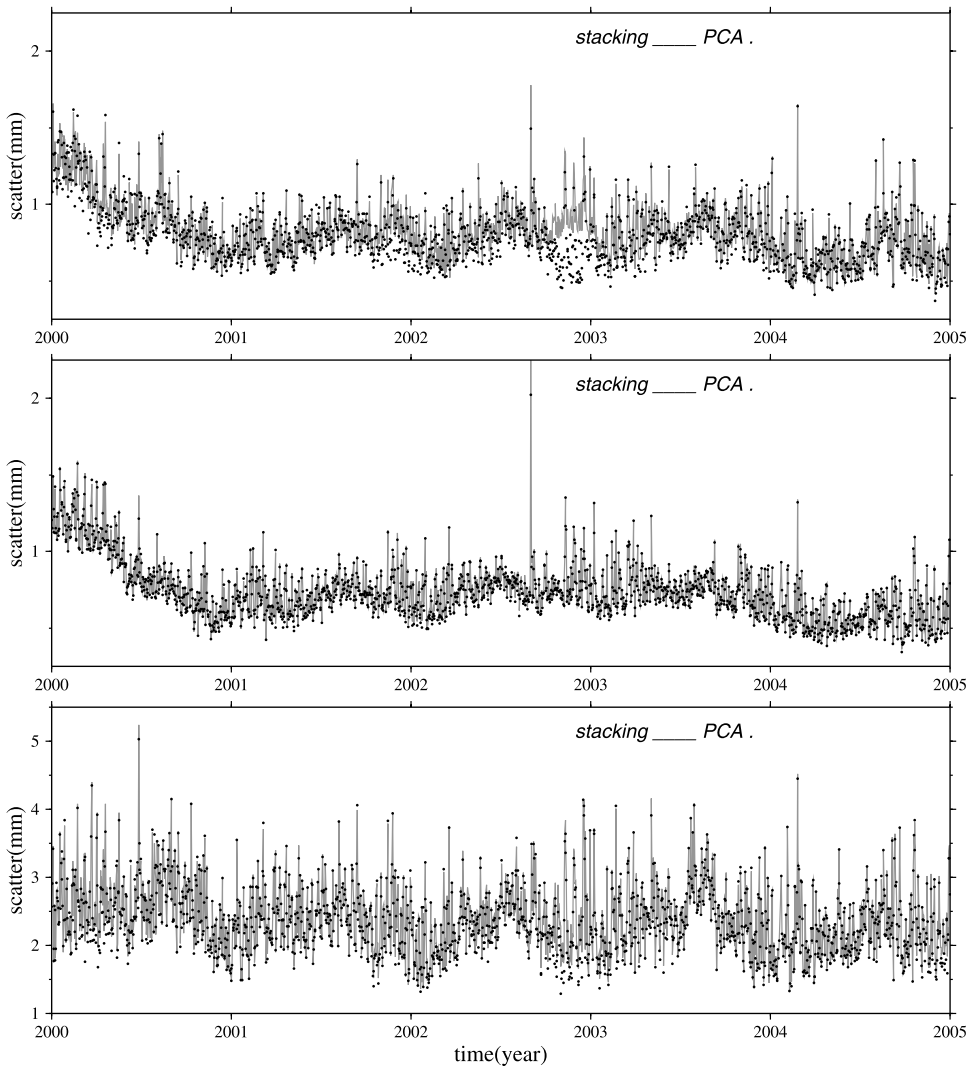

Figure 11. Comparison of the daily scatters of the regionally filtered time series by the stacking approach (solid line) and the PCA approach (dots). The scatters are calculated using L1 norm (see text for details). (top) East component. (middle) North component. (bottom) Vertical component.

temperature-dependent, such as small-scale atmospheric perturbations, antenna thermal noise, and thermoelastic strain [*Prawirodirdjo et al.*, 2006].

5. Conclusions

[37] The proper characterization and suppression of GPS errors is critical in maximizing the accuracy achievable from regional networks. Using the combination of PCA and KLE (PCA/KLE), we have reestablished that spatiotemporal correlated errors are the dominant error source in daily GPS positions by examining the time series of 148 SCIGN stations over a 5-year period. We have demonstrated that PCA/KLE provides a more general approach to spatial filtering by being able to identify the principal components in the spatial distribution of the CME. Our analysis shows that the temporal behavior of the CME is not purely random. The spatial characteristics of the CME are close to uniform for all east, north and vertical components, which implies a very long wavelength for the CME, compared to the spatial extent of SCIGN, and also explains the good agreement of PCA/KLE with the traditional stacking approach [*Wdowinski et al.*, 1997]. PCA/KLE can also be used to investigate spatiotemporally correlated geophysical signals in the position time series, such as mass loading caused by seasonal deformation [*Tiampo et al.*, 2004], but is beyond the scope of this paper.

[38] Although the assumption of spatial uniform distribution of CME is successful for SCIGN and several other similar size regional networks, it is unclear what the spatial

Table 1. Scatter Comparison[a]

Component	Unfiltered	Stacking	PCA	KLE
East (L1)	2.46	0.83	0.80	0.82
North (L1)	1.65	0.75	0.74	0.75
Vertical (L1)	5.02	2.40	2.37	2.38
East (L2)	3.30	1.19	1.16	1.18
North (L2)	2.12	1.11	1.10	1.12
Vertical (L2)	6.44	3.23	3.19	3.21

[a]L1 norm scatters (in mm) and L2 norm scatters for the residual coordinate time series from 2000 to 2004 at 148 CGPS stations in southern California. The L1 and L2 norm scatters are defined in the text.

limit of this assumption is. To extend regional filtering to larger networks, it is critical to understand the nature and the wavelength of the CME. In particular, different software and different analysis strategies could lead to different systematic errors, generating different spatiotemporal CME patterns and wavelengths. Using a continental-scale GPS network (North America and Mexico) data, *Márquez-Azúa and Demets* [2003] found that the spatial cross correlations of the residual time series are high within 1000 km distance, then gradually decrease to zero at about 6000 km distance. Existing larger-scale CGPS networks such as the Japanese GEONET [*Miyazaki et al.*, 1997] and new networks, in particular the in-construction CGPS network of EarthScope's Plate Boundary Observatory (PBO) will allow us to more fully understand the source of regional CME. The spatiotemporal patterns exhibited in our analysis suggest that these CME are likely caused by unmodeled or mismodeled motions of satellite orbits, reference frame or EOP [*Wdowinski et al.*, 1997]. Large-scale atmosphere effects, receiver and satellite antenna phase center mismodeling are also potential candidates for the CME.

[39] Other outstanding questions are the nature of the seasonal components remaining in the filtered CGPS position time series, and their classification as signal or noise. When constructing the residual coordinate time series, we implicitly assume that all secular and seasonal components are signals. However, previous studies [e.g., *Dong et al.*, 2002; *Kedar et al.*, 2003] have shown that some of the seasonal effects are due to systematic errors in the daily GPS solution time series.

[40] Finally, implicit to the analysis presented here is that CME analysis can be applied to any size network, regardless of it size. The idea to perform subnetwork CME analysis based on the characteristics of the PC error analysis of the entire network should be further explored. How to define and quantify the CME is still an open question.

[41] **Acknowledgments.** We are grateful to K. F. Tiampo, J. W. Parker, and K. Stark for helpful discussions. We would like to thank Shimon Wdowinski and an anonymous reviewer for their constructive comments, which improve the manuscript substantially. The maps in this paper were generated using the public domain Generic Mapping Tools (GMT) software [*Wessel and Smith*, 1995]. This research described in this paper was supported under a contract with NASA at Jet Propulsion Laboratory, California Institute of Technology, and at Scripps Institution of Oceanography.

References

Altamimi, Z., P. Sillard, and C. Boucher (2002), ITRF2000: A new release of the International Terrestrial Reference Frame for earth science applications, *J. Geophys. Res.*, *107*(B10), 2214, doi:10.1029/2001JB000561.

Argus, D. F., M. B. Heflin, G. Peltzer, F. Crampe, and F. H. Webb (2005), Interseismic strain accumulation and anthropogenic motion in metropolitan Los Angeles, *J. Geophys. Res.*, *110*, B04401 doi:10.1029/2003JB002934.

Bawden, G. W., W. Thatcher, R. S. Stein, K. W. Hudnut, and G. Peltzer (2001), Tectonic contraction across Los Angeles after removal of groundwater pumping effects, *Nature*, *412*, 812–815.

Beavan, J., P. Tregoning, M. Bevis, T. Kato, and C. Meertens (2002), Motion and rigidity of the Pacific plate and implications for plate boundary deformation, *J. Geophys. Res.*, *107*(B10), 2261, doi:10.1029/2001JB000282.

Blewitt, G., D. Lavallee, P. Clarke, and K. Nurutdinov (2001), A new global mode of Earth deformation, Seasonal cycle detected, *Science*, *294*, 2342–2345.

Bock, Y., et al. (1993), Detection of crustal deformation from Landers earthquake sequence using continuous geodetic measurements, *Nature*, *361*, 337–340.

Bock, Y., L. Prawirodirdjo, and T. I. Melbourne (2004), Detection of arbitrarily large dynamic ground motions with a dense high-rate GPS network, *Geophys. Res. Lett.*, *31*, L06604, doi:10.1029/2003GL019150.

Calais, E. (1999), Continuous GPS measurements across the western Alps, 1996–1998, *Geophys. J. Int.*, *138*, 221–230.

Dong, D., T. A. Herring, and R. W. King (1998), Estimating regional deformation from a combination of space and terrestrial geodetic data, *J. Geod.*, *72*, 200–214.

Dong, D., P. Fang, Y. Bock, M. K. Cheng, and S. Miyazaki (2002), Anatomy of apparent seasonal variations from GPS-derived site position time series, *J. Geophys. Res.*, *107*(B4), 2075, doi:10.1029/2001JB000573.

Dong, D., T. Yunck, and M. Heflin (2003), Origin of the International Terrestrial Reference Frame, *J. Geophys. Res.*, *108*(B4), 2200, doi:10.1029/2002JB002035.

Dragert, H., K. Wang, and T. James (2001), A silent slip event on the deeper Cascadia subduction interface, *Science*, *292*, 1525–1528.

Farrell, W. E. (1972), Deformation of the Earth by surface loads, *Rev. Geophys.*, *10*, 761–797.

Fialko, Y. (2004), Evidence of fluid-filled upper crust from observations of postseismic deformation due to the 1992 M_w7.3 Landers earthquake, *J. Geophys. Res.*, *109*, B08401, doi:10.1029/2004JB002985.

Fukunaga, K. (1990), *Introduction to Statistical Pattern Recognition*, Elsevier, New York.

Hammond, W. C., and W. Thatcher (2004), Contemporary tectonic deformation of the Basin and Range province, western United States: 10 years of observation with the Global Positioning System, *J. Geophys. Res.*, *109*, B08403, doi:10.1029/2003JB002746.

Hatanaka, Y., M. Sawada, A. Horita, M. Kusaka, J. Johnson, and C. Rocken (2001), Calibration of antenna-radome and monument-multipath effect of GEONET–part 2: Evaluation of the phase map by GEONET data, *Earth Planets Space*, *53*, 23–30.

Herring, T. A. (2005), GLOBK: Global Kalman filter VLBI and GPS analysis program version 10.01, Mass. Inst. of Technol., Cambridge.

Hudnut, K. W., et al. (1996), Co-seismic displacements of the 1994 Northridge, California, earthquake, *Bull. Seismol. Soc. Am.*, *86*(1B), s19–s36.

Hudnut, K. W., N. King, J. Galetzka, K. Stark, J. Behr, A. Aspiotes, S. van Wyk, R. Moffitt, S. Dockter, and F. Wyatt (2002), Continuous GPS observations of postseismic deformation following the Oct. 16, 1999 Hector Mine earthquake (M_w7.1), *Bull. Seismol. Soc. Am.*, *92*, 1400–1422.

Johansson, J. M., et al. (2002), Continuous GPS measurements of postglacial adjustment in Fennoscandia: 1. Geodetic results, *J. Geophys. Res.*, *107*(B8), 2157, doi:10.1029/2001JB000400.

Kedar, S., G. A. Hajj, B. D. Wilson, and M. B. Heflin (2003), The effect of the second order GPS ionospheric correction on receiver positions, *Geophys. Res. Lett.*, *30*(16), 1829, doi:10.1029/2003GL017639.

King, R. W., and Y. Bock (2005), Documentation for the GAMIT GPS analysis software version 10.20, Mass. Inst. of Technol., Cambridge.

Larson, K., J. Freymueller, and S. Philipsen (1997), Global plate velocities from the Global Positioning System, *J. Geophys. Res.*, *102*, 9961–9982.

Larson, K., P. Bodin, and J. Gomberg (2003), Using 1-Hz GPS data to measure deformations caused by the Denali Fault earthquake, *Science*, *300*, 1421–1424.

Mangiarotti, S., A. Cazenave, L. Soudarin, and J. F. Cretaux (2001), Annual vertical crustal motions predicted from surface mass redistribution and observed by space geodesy, *J. Geophys. Res.*, *106*, 4277–4291.

Mao, A., C. G. Harrison, and T. H. Dixon (1999), Noise in GPS time series, *J. Geophys. Res.*, *104*, 2797–2816.

Márquez-Azúa, B., and C. DeMets (2003), Crustal velocity field of Mexico from continuous GPS measurements, 1993 to June 2001: Implications for the neotectonics of Mexico, *J. Geophys. Res.*, *108*(B9), 2450, doi:10.1029/2002JB002241.

Melbourne, T. I., and F. H. Webb (2002), Precursory transient slip during the 2001 M_w = 8.4 Peru earthquake sequence from continuous GPS, *Geophys. Res. Lett.*, *29*(21), 2032, doi:10.1029/2002GL015533.

Melbourne, T. I., F. H. Webb, J. M. Stock, and C. Reigber (2002), Rapid postseismic transients in subduction zones from continuous GPS, *J. Geophys. Res.*, *107*(B10), 2241, doi:10.1029/2001JB000555.

Menke, W. (1984), *Geophysical Data Analysis, Discrete Inverse Theory*, Elsevier, New York.

Miller, M. M., D. J. Johnson, T. H. Dixon, and R. K. Dokka (2001a), Refined kinematics of the Eastern California shear zone from GPS observations, 1993–1998, *J. Geophys. Res.*, *106*, 2245–2263.

Miller, M. M., D. J. Johnson, C. M. Robin, H. Dragert, K. Wang, A. Qamar, and C. Goldfinger (2001b), GPS-determination of along-strike variation in Cascadia margin kinematics: Implication for relative plate motion, subduction zone coupling, and permanent deformation, *Tectonics*, *20*, 161–176.

Miller, M. M., T. Melbourne, D. J. Johnson, and W. Q. Sumner (2002), Periodic slow earthquakes from the Cascadia subduction zone, *Science*, *295*, 2423.

Miyazaki, S., and K. Heki (2001), Crustal velocity field of southwest Japan: Subduction and arc-arc-collision, *J. Geophys. Res.*, *106*, 4305–4326.

Miyazaki, S., T. Saito, M. Sasaki, Y. Hatanaka, and Y. Iimura (1997), Expansion of GSI's nationwide GPS array, *Bull. Geogr. Surv. Inst.*, *43*, 23–34.

Miyazaki, S., T. Iwabuchi, K. Heki, and I. Naito (2003), An impact of estimating tropospheric delay gradients on precise positioning in the summer using the Japanese nationwide GPS array, *J. Geophys. Res.*, *108*(B7), 2335, doi:10.1029/2000JB000113.

Murray, M. H., W. H. Prescott, R. Bürgmann, J. T. Freymueller, P. Segall, J. Svarc, S. D. J. Williams, M. Lisowski, and B. Romanowicz (1998), The deformation field of the Pacific-North America plate boundary zone in northern California from geodetic data, 1973–1989, *Eos Trans. AGU*, *79*(45), Fall Meeting Suppl., F192.

Nikolaidis, R. (2002), Observation of geodetic and seismic deformation with the Global Positioning System, Ph.D. thesis, Univ. of Calif., San Diego.

Ohtani, R., P. Segall, and J. McGuire (2004), The Network Strain Filter—A new tool for monitoring and detecting transient deformation signals in GPS arrays, *Eos Trans. AGU*, *85*(47), Fall Meet. Suppl., Abstract G21A-0136.

Parker, J. W. (2001), Analysis and modeling of southern California deformation, in *APEC Cooperation for Earthquake Simulation (ACES), 2nd ACES Workshop Proceedings*, Univ. of Queensland, Brisbane, Australia.

Penna, N. T., and M. P. Stewart (2003), Aliased tidal signatures in continuous GPS height time series, *Geophys. Res. Lett.*, *30*(23), 2184, doi:10.1029/2003GL018828.

Pollitz, F. F., G. Peltzer, and R. Burgmann (2000), Mobility of continental mantle: Evidence from postseismic geodetic observations following the 1992 Landers earthquake, *J. Geophys. Res.*, *105*, 8035–8054.

Pollitz, F. F., C. Wicks, and W. Thatcher (2001), Mantle flow beneath a continental strike-slip fault, Postseismic deformation after the 1999 Hector Mine earthquake, *Science*, *293*, 1814–1818.

Prawirodirdjo, L., and Y. Bock (2004), Instantaneous global plate motion mode from 12 years of continuous GPS observations, *J. Geophys. Res.*, *109*, B08405, doi:10.1029/2003JB002944.

Prawirodirdjo, L., Y. Ben-Zion, and Y. Bock (2006), Observation and modeling of thermoelastic strain in Southern California Integrated GPS Network daily position time series, *J. Geophys. Res.*, *111*, B02408, doi:10.1029/2005JB003716.

Preisendorfer, R. W. (1988), *Principal Component Analysis in Meteorology and Oceanography*, edited by C. D. Mobley, *Develop. Atmos. Sci.*, vol. 17, Elsevier, New York.

Rundle, J. B., W. Klein, K. Tiampo, and S. Gross (2000), Linear pattern dynamics in nonlinear threshold systems, *Phys. Rev. E*, *61*, 2418–2431.

Savage, J. (1988), Principal component analysis of geodetically measured deformation in Long Valley caldera, eastern California, *J. Geophys. Res.*, *93*, 1983–1987.

Savage, J. (1995), Principal component analysis of interseismic deformation in southern California, *J. Geophys. Res.*, *100*, 12,691–12,701.

Scherneck, H. G. (1991), A parameterized solid Earth tide model and ocean tide loading effects for global geodetic baseline measurements, *Geophys. J. Int.*, *106*, 677–694.

Scherneck, H.-G., J. M. Johansson, and R. Haas (2000), BIFROST project: Studies of variations of absolute sea level in conjunction with postglacial rebound in Fennoscandia, in *Towards an Integrated Global Geodetic Observing System (IGGOS)*, *Int. Assoc. Geod. Symp.*, vol. 120, edited by R. Rammel et al., pp. 241–244, Springer, New York.

Sella, G. F., T. H. Dixon, and A. L. Mao (2002), REVEL: A model for recent plate velocities from space geodesy, *J. Geophys. Res.*, *107*(B4), 2081, doi:10.1029/2000JB000033.

Shen, Z. K., et al. (1997), Crustal deformation measured in southern California, *Eos Trans. AGU*, *78*(43), 477, 482.

Simons, M., Y. Fialko, and L. Rivera (2002), Coseismic deformation from the 1999 M_w 7.1 Hector Mine, California, Earthquake as Inferred from InSAR and GPS Observations, *Bull. Seismol. Soc. Am.*, *92*, 1390–1402.

Smith, K. D., D. von Seggern, G. Blewitt, L. Preston, J. G. Anderson, B. P. Wernicke, and J. L. Davis (2004), Evidence for deep magma injection beneath Lake Tahoe, Nevada-California, *Science*, *305*, 1277–1280.

Tiampo, K. F., J. B. Rundle, W. Klein, Y. Ben-Zion, and S. McGinnis (2004), Using eigenpattern analysis to constrain seasonal signals in southern California, *Pure Appl. Geophys.*, *161*, 1991–2003.

vanDam, T. M., and J. M. Wahr (1987), Displacements of the Earth's surface due to atmospheric loading: Effects on gravity and baseline measurements, *J. Geophys. Res.*, *92*, 1281–1286.

vanDam, T. M., G. Blewitt, and M. B. Heflin (1994), Atmospheric pressure loading effects on Global Positioning System coordinate determinations, *J. Geophys. Res.*, *99*, 23,939–23,950.

vanDam, T. M., J. M. Wahr, P. C. D. Milly, A. B. Shmakin, G. Blewitt, D. Lavallee, and K. M. Larson (2001), Crustal displacements due to continental water loading, *Geophys. Res. Lett.*, *28*, 651–654.

Watson, K. M., Y. Bock, and D.T. Sandwell (2002), Satellite interferometric observations of displacements associated with seasonal groundwater in the Los Angeles basin, *J. Geophys. Res.*, *107*(B4), 2074, doi:10.1029/2001JB000470.

Wdowinski, S., Y. Bock, J. Zhang, P. Fang, and J. Genrich (1997), Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake, *J. Geophys. Res.*, *102*, 18,057–18,070.

Wdowinski, S., Y. Bock, G. Baer, L. Prawirodirdjo, N. Bechor, S. Naaman, R. Knafo, Y. Forrai, and Y. Melzer (2004), GPS measurements of current crustal movement along the Dead Sea Fault, *J. Geophys. Res.*, *109*, B05403, doi:10.1029/2003JB002640.

Wessel, P., and W. H. F. Smith (1995), New version of the Generic Mapping Tools released, *Eos Trans. AGU*, *76*, 329.

Williams, S. D. P., Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson (2004), Error analysis of continuous GPS position time series, *J. Geophys. Res.*, *109*, B03412, doi:10.1029/2003JB002741.

Zhang, J. (1996), Continuous GPS measurements of crustal deformation in southern California, Ph.D. dissertation, Univ. of Calif., San Diego.

Zhang, J., Y. Bock, H. Johnson, P. Fang, S. Williams, J. Genrich, S. Wdowinski, and J. Behr (1997), Southern California Permanent GPS Geodetic Array: Error analysis of daily position estimates and site velocities, *J. Geophys. Res.*, *102*, 18,035–18,055.

Y. Bock, P. Fang, P. Jamason, and L. Prawirodirdjo, Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, La Jolla, CA 92093, USA. (ybock@ucsd.edu; pfang@pgga.ucsd.edu; pjamason@gpsmail.ucsd.edu; linette@gpsmail.ucsd.edu)

D. Dong, S. Kedar, and F. Webb, Jet Propulsion Laboratory, MS 238-600, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA. (danan.dong@jpl.nasa.gov; sharon.kedar@jpl.nasa.gov; frank.h.webb@jpl.nasa.gov)